Claims

1. A pseudo-sequence method for comparing a first 7TM receptor with one or more further 7TM receptors with respect to the physicochemical properties of selected amino acid residues of their binding sites, the method comprising the steps of:

i) optionally, aligning part of or all of the amino acid sequence of the first 7TM receptor with part of or all of the amino acid sequence of the one or more further 7TM receptors,

ii) selecting, in a sequential or non-sequential order, at the most 12 amino acid residues per helix and/or extracellular loops, which are involved in one or more binding sites of each 7TM receptor,

iii) forming a pseudo-sequence comprising at the most 50 amino acid residues from the selected sequential or non-sequential amino acid residues,

iv) for each 7TM receptor assigning one or more physicochemical descriptors to the amino acid residues of the selected amino acid pseudo-sequence involved in one or more binding sites,

v) optionally, for each 7TM receptor mathematically manipulating the physicochemical descriptors of step iv) to obtain a simplified measure of the physicochemical properties of the binding site,

vi) for each 7TM receptor generating a similarity score as defined herein by comparing the physicochemical descriptor or, if relevant, the simplified measure for the first 7TM receptor with the physicochemical descriptors or, if relevant, the simplified measures for the one or further 7TM receptors,

vii) optionally, ranking the 7TM receptors with respect to the physicochemical properties of their binding sites according to the similarity scores obtained in step vi).

2. A method according to claim 1, wherein the comparison is made without using data related to binding affinity of a ligand to a 7TM receptor.

3. A method according to claim 1 for classifying 7TM receptors according to the physicochemical properties of their binding sites.

4. A method according to claim 3, wherein the classification is made without using data related to binding affinity of a ligand to a 7TM receptor.

5. A method according to any of the preceding claims, wherein step ii) as defined in claim 1 comprises selecting, in a sequential or non-sequential order, at the most 11 such as, e.g., at the most 10, at the most 9, at the most 8, at the most 7 or at the most 6 amino acid residues per helix and/or extracellular loops, which are involved in one or more binding sites of each 7TM receptor.

6. A method according to any of the preceding claims, wherein step iii) as defined in claim 1 comprises forming a pseudo-sequence comprising at the most 50 such as, e.g., at the most 45, at the most 40, at the most 35 or at the most 30 amino acid residues from the selected sequential or non-sequential amino acid residues.

7. A drug discovery method for identifying ligands, which bind to a first 7TM receptor and potentially bind to one or more further 7TM receptors, the method comprising the steps of i) to vii) as defined in claim 1, 5 or 6 and the further steps of

viii) selecting from one to about 100 further 7TM receptors which have the closest similarity scores to the first 7TM receptor,

ix) identifying ligands which potentially bind to those further 7TM receptors selected in step vii) by selecting ligands that bind to the first 7TM receptor.

8. A drug discovery method for identifying ligands which bind to a first 7TM receptor and to one or more further 7TM receptors, the method comprising the steps of i) to vii) as defined in claim 1, 5 or 6 and the further steps of:

viii) selecting from one to about 100 further 7TM receptors which have the closest similarity scores to the first 7TM receptor,

ix) screening ligands that bind to the first 7TM receptor against the selected 7TM receptors of step viii).

9. A drug discovery method for identifying a potential lead compound for a first 7TM

receptor, the method comprising the steps of i) to vii) as defined in claim 1, 5 or 6 and the further steps of

viii) selecting from one to about 100 further 7TM receptors which have the closest similarity scores to the first 7TM receptor,

ix) identifying ligands that bind to said one or more further 7TM receptors to construct a library including a potential lead compound for the first 7TM receptor.

10. A drug discovery method for identifying a lead compound for a first 7TM receptor, the method comprising the steps of i) to vii) as defined in claim 1, 5 or 6 and the further steps of

viii) selecting from one to about 100 further 7TM receptors which have the closest similarity scores to the first 7TM receptor,

ix) identifying ligands that bind to said one or more further 7TM receptors to construct a library, and

x) screening said library against the first 7TM receptor to identify a lead compound for the first 7TM receptor.

11. A drug discovery method for constructing a pharmacophore model for a first 7TM receptor, the method comprising the steps of i) to vii) as defined in claim 1, 5 or 6 and the further steps of

viii) selecting from one to about 100 further 7TM receptors which have the closest similarity scores to the first 7TM receptor,

ix) identifying ligands that bind to said one or more further 7TM receptors to construct a pharmacophore model.

12. A drug discovery method according to claim 10, wherein the first 7TM receptor is one for which no ligands have been identified.

13. A drug discovery method according to claim 10 or 11, wherein the first 7TM receptor is an orphan receptor.

14. A method according to any of claims 7-12, wherein from one to 50 further 7TM receptors is/are selected in step viii).

15. A method according to any of claims 7-12, wherein from one to 25 further 7TM receptors is/are selected in step viii).

16. A method according to any of claims 7-12, wherein from one to 15 further 7TM receptors is/are selected in step viii).

17. A method according to any of the preceding claims, wherein the method is executed by a computer under the control of a program and the computer includes a memory for storing said program.

18. A method according to any of the preceding claims, wherein step i) is included and the alignment is based on a model developed for 7TM receptors.

19. A method according to claim 18, wherein the 7TM receptors are Class A, Class B, Class C or taste receptors.

20. A method according to any of the preceding claims, wherein step i) is included and the alignment is made with respect to transmembrane positioning of α-helices of 7TM receptors.

21. A method according to any of the preceding claims, wherein the binding site includes amino acid residues located in one or more extracellular loops of the 7TM receptors.

22. A method according to any of the preceding claims, wherein the binding site includes amino acid residues located in one or more subsites of the binding site and in one or more extracellular loops of the 7TM receptors.

23. A method according to any of the preceding claims, wherein the physicochemical descriptors reflect 7TM receptor-ligand interaction features of the amino acid residues.

24. A method according to any of the preceding claims, wherein the physicochemical descriptors are chosen to reflect hydrophobic, electronic, steric, hydrogen bonding or other properties of the amino acid residues.

25. A method according to any of the preceding claims, wherein the physicochemical descriptors reflect 3-dimensional features of the amino acid residues.

26. A method according to any of the preceding claims, wherein the physicochemical descriptors are selected from descriptors used in quantitative structure-activity relationships (QSAR), Principle Component Regression (PCR) and Partial Least-Squares (PLS) analysis of peptides.

27. A method according to any of claims 23-26, wherein the physicochemical descriptors are selected from molecular weight (MW), van der Waals volume, van der Waals radius, molar refractivity (MR), STERIMOL parameters (L, B_1, B_5), Parachor (P_r), polar surface area, non-polar surface area, total surface area, ionisation constant (pK_{COOH}, pK_{NH2}), isoelectric point, net charge at pH 7, partition coefficient (log P), calculated partition coefficient (clog P, Prolog P, Maclog P), distribution coefficient (log D), TLC retention time, HPLC retention time, HPLC capacity factor log k, ^{1}H NMR chemical shift, ^{13}C NMR chemical shift, steric and electrostatic 3D-property MS-WHIM indexes, calculated interaction energies, isotropic surface area (ISA), electronic charge index (ECI), charge transfer for carbons (CT), Lewis basicity (LB), Lewis acidity (LA), maximum electrostatic potential (V_{max}), minimum electrostatic potential (V_{min}), maximum local ionization energy (I_{max}), minimum local ionization energy (I_{min}), conformational strain energy (ΔH_{strain}), molecular electrostatic potential (MEP) on Connolly molecular surface, local flexibility (Fr), flexibility index (Fb), chain flexibility (FO), occupied volume by a residue buried in globular protein, bulkiness defined as the ratio of the side-chain volume to its length, total energy (E_{total}), heat of formation (ΔH_f), energy of highest occupied molecular orbital (E_{HOMO}), energy of lowest unoccupied molecular orbital (E_{LUMO}), dipole moment (μ), polarizability (α), most positive partial charge on a hydrogen atom (qH+), most negative partial charge in the molecule (q−), partial charges on the oxygen and carbon atoms (qC, qO) of the carbonyl group, integrated molecular transform (FTm), integrated electronic transform (FTe), Integrated charge transform (FTc), normalized molecular moment (Mn), electronic moment (Me), charge moment (Mc), absolute electronegativity (EN), absolute hardness (HA).

28. A method according to any of claims 17-25, wherein the physicochemical descriptors include indicator variables such as, e.g., 1 and 0.

29. A method according to claim 28, wherein the indicator variables denote absence or presence of aromatic side chains, hydrophobic side chains, negatively charged side chains, positively charged side chains, polar side chains, hydrogen-bond donating side chains, hydrogen-bond accepting side chains and/or other selected features.

30. A method according to any of the preceding claims including step v), wherein the physicochemical descriptors are weighted in step v).

31. A method according to any of the preceding claims including step v), wherein a simplified measure of the physicochemical properties of the binding site is obtained from principal component analysis (PCA) of the physicochemical descriptors.

32. A method according to any of the preceding claims, wherein the generation of a similarity score in step vi) is based upon a pattern recognition method.

33. A method according to any of the preceding claims, wherein the generation of the similarity score involves a Principal Component Analysis (PCA) reducing the number of descriptors to a few principal components.

34. A method according to any of the preceding claims, wherein the generation of the similarity score in step vi) is based upon Euclidian Distance Measure: $d(F1,F2) = \text{sqrt}(F1 - F2)^2$.

35. A method according to claim 28 or 29, wherein the generation of the similarity score in step v) is based upon a Tanimoto Similarity Measure: $TC = BC / (B1 + B2 - BC)$.

36. A method according to claim 28 or 29, wherein the generation of the similarity score in step v) is based upon a Tversky Similarity Measure: $TC = BC / (\alpha * B1Unique + \beta * B2Unique + BC)$, wherein α are prototype features and β variant features.

37. A method according to any of the preceding claims, wherein the step vii) is included.

38. A method according to any of the preceding claims, wherein the similarity score or, if relevant, the ranking is based upon a 2- or 3-dimensional graphical representation.

39. Use of a pharmacophore according to claim 11 for *in silico* screening.

40. Use of a pharmacophore according to claim 11 for construction of a library.

41. Use of a pharmacophore according to claim 11 for design of a ligand.

42. Use of a method according to any of claims 1-38 to identify receptors, which are likely to cause a selectivity problem during drug development of a drug interacting with a given receptor.

43. Use of a method according to any of claims 1-38 to identify differences in subsites of binding sites between 7TM receptors as means to improve receptor selectivity of a drug towards a given 7TM receptor.